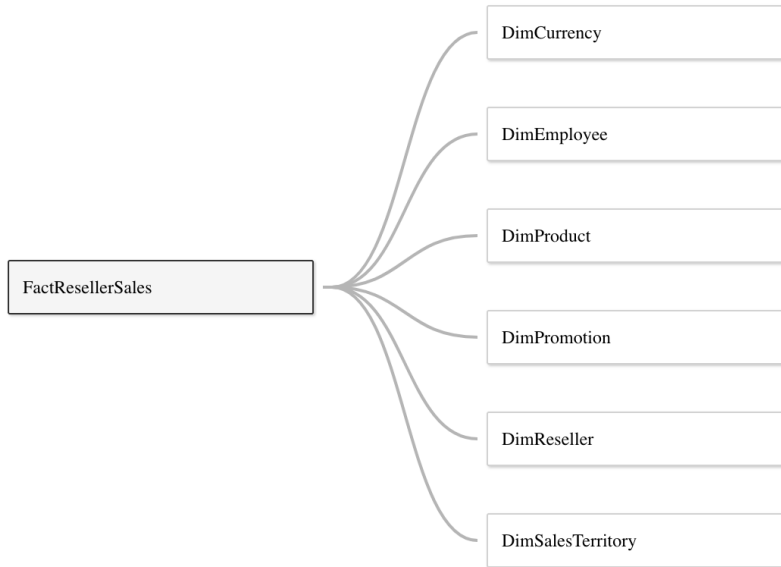


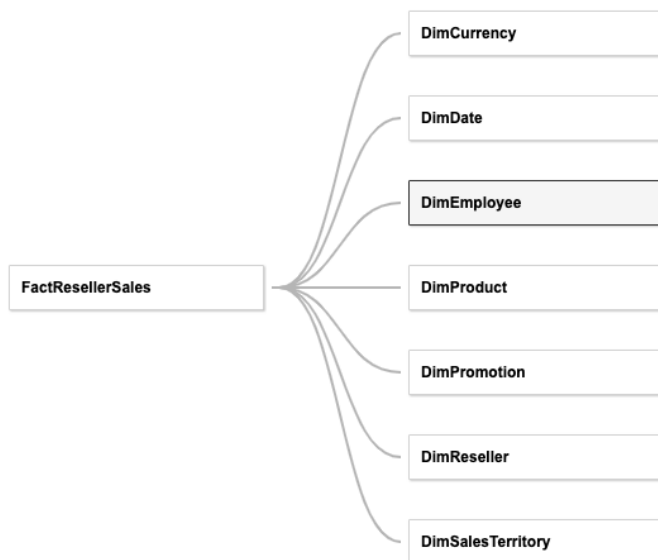
Tableau Analysis

This assignment adopt the sample Adventure Works Data Warehouse, connecting the tables FactResellerSales, DimCurrency, DimEmployee, DimProduct, DimPromotion, DimSeller, DimSalesTerritory.



Connected tables snapshot: -

FactResellerSales+ (AdventureWorksDW2017)



1. Cross Table Calculation and Aggregation:

(1). Create and display a cross table with **“Business Type”** firstly (from DimReseller) and **“Sales Territory Region”** secondly (from DimSalesTerritory) as rows, with **Year of “Order Date”** and **Quarter of “Order Date”** (discrete types) from FactResellerSales as columns; The value of the cells in this cross table is the **average of “Sales Amount”** from FactResellerSales;

Besides the values in the cells, generate and display **row grand totals**, **column grand totals** and **all subtotals** in the same cross table;

Pages

Columns

YEAR(Order Date)

QUARTER(Order D..

Rows

Business Type

Sales Territory Region

Filters

Sheet 1

Marks

Automatic

ColorSizeText

DetailTooltip

AVG(Sales Am...

		2011			2012					2013					Grand Total
Business Ty..	Sales Territ..	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total	Grand Total
Specialty Bike Shop	Southeast	06	800	1,847	733	1,027	894	740	815	438	844	932	859	698	979
	Southwest	75	731	1,000	594	765	924	619	676	527	821	978	606	694	739
	United King..		886	886	848	765	835	528	740	524	1,115	932	982	820	790
	Total	35	859	1,504	660	904	926	643	739	464	842	1,019	706	691	846
Value Added Reseller	Australia							663	663	696	786	1,021	820	811	805
	Canada	03	1,266	1,587	1,236	1,361	1,989	1,241	1,396	977	1,116	1,294	995	1,079	1,276
	Central	01	2,030	1,788	1,330	1,441	1,970	1,418	1,496	1,262	1,304	1,671	1,071	1,305	1,466
	France		1,039	1,039	793	1,184	964	563	811	1,254	1,706	1,711	1,956	1,603	1,386
	Germany							1,102	1,102	814	1,084	1,195	896	983	990
	Northeast	91	1,346	1,299	1,384	1,757	1,796	1,107	1,492	900	1,001	1,075	932	976	1,263
	Northwest	85	1,806	2,359	1,708	1,936	1,941	1,684	1,799	1,375	1,526	1,948	1,354	1,529	1,782
	Southeast	90	1,032	1,728	1,319	1,448	1,508	1,108	1,326	1,031	1,000	1,089	1,302	1,085	1,322
	Southwest	80	1,932	1,876	1,421	1,604	1,898	1,466	1,566	1,168	1,180	1,499	1,162	1,240	1,479
	United King..				1,118	1,251	1,828	1,045	1,240	1,019	1,193	1,347	1,201	1,172	1,198
	Total	52	1,588	1,798	1,359	1,536	1,830	1,300	1,473	1,086	1,191	1,421	1,138	1,195	1,391
Warehouse	Australia							2,289	2,289	1,669	1,616	974	1,255	1,402	1,404
	Canada	01	1,770	1,828	1,327	1,151	1,156	1,283	1,236	1,356	1,247	1,205	1,116	1,248	1,369
	Central	99	1,988	2,408	1,121	996	936	1,216	1,074	1,173	1,168	945	1,216	1,128	1,389
	France		1,010	1,010	1,742	1,289	1,053	1,320	1,376	1,492	1,472	1,179	1,435	1,411	1,391
	Germany							1,323	1,323	1,473	1,052	1,180	1,122	1,224	1,234
	Northeast	89	1,352	1,285	1,411	1,235	1,158	1,159	1,248	1,342	1,013	903	1,157	1,122	1,207
	Northwest	24	2,458	3,047	1,209	1,158	1,028	1,213	1,161	1,599	1,435	1,345	1,413	1,461	1,556
	Southeast	78	2,217	2,484	1,243	1,120	737	1,054	1,046	1,178	993	938	1,339	1,109	1,499
	Southwest	39	1,397	1,779	1,546	1,181	1,076	1,484	1,345	1,512	1,307	1,351	1,172	1,359	1,428
	United King..		1,212	1,212	1,457	1,196	876	1,010	1,132	1,593	1,380	1,178	1,821	1,477	1,343
	Total	06	1,736	2,005	1,410	1,174	1,054	1,287	1,245	1,441	1,274	1,178	1,293	1,309	1,396
Grand Total		72	1,550	1,851	1,264	1,296	1,343	1,217	1,274	1,131	1,183	1,264	1,150	1,176	1,322

(2). What is the subtotal of sales amount from **Value Added Reseller** in **2011 Q3**?

Answer: The subtotal of average of sales amount from Value Added Reseller in 2011 Q3 is **2,552**.

Sheet 1

Business Ty..	Sales Territ..	2010	2011				Order Date				
		Total	Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4
Specialty Bike Shop	Northwest	1,154	2,501	3,092	3,042	968	2,069	574	854	914	611
	Southeast	2,214	2,454	2,115	2,806	800	1,847	733	1,027	894	740
	Southwest	647	706	1,545	1,175	731	1,000	594	765	924	619
	United King..										
	Total	1,261	1,806	1,856	2,035	859	1,504	660	904	926	643
Value Added Reseller	Australia										663
	Canada	928	1,632	1,186	2,203	1,266	1,587	1,236	1,361	1,989	1,241
	Central	635	1,203	1,969	1,701	2,030	1,788	1,330	1,441	1,970	1,418
	France					1,039	1,039	793	1,184	964	563
	Germany										1,102
	Northwest	748	1,024	1,322	1,391	1,346	1,299	1,384	1,757	1,796	1,107
	Northwest	2,136	2,361	2,294	3,185	1,806	2,359	1,708	1,936	1,941	1,684
	Southwest	1,602	1,676	2,246	2,390	1,032	1,728	1,319	1,448	1,508	1,108
	Southwest	623	1,640	1,562	2,280	1,932	1,876	1,421	1,664	1,898	1,466
	United King..										1,045
	Total	1,310	1,623	1,771	2,252						1,300
Warehouse	Australia										2,289
	Canada	1,477	1,785	2,076	1,801						1,283
	Central		2,944	2,903	2,399						1,216
	France										1,320
	Germany										1,323
	Northwest	629	1,279	1,408	989	1,302	1,289	1,311	1,276	1,138	1,159
	Northwest	715	3,281	3,991	2,724	2,458	3,047	1,209	1,158	1,028	1,213
	Southwest	1,833	2,548	3,082	2,278	2,217	2,484	1,243	1,120	737	1,054
	Southwest	1,410	1,924	2,409	1,739	1,397	1,779	1,546	1,181	1,076	1,484
	United King..										1,010
	Total					1,212	1,212	1,457	1,196	876	1,010

(3). Based on the prior cross table, generate and display **quick table calculation** with **percentage of total**, using **pane (down)**;

Pages

Columns: YEAR(Order Date) QUARTER(Order Date)

Rows: Business Type Sales Territory Region

Filters

Sheet 1

Business Ty..	Sales Territ..	2011				2012					2013				Grand Total
		Q3	Q4	Total		Q1	Q2	Q3	Q4	Total	Q1	Q2	Q3	Q4	Total
Specialty Bike Shop	Southwest	9%	93.2%	122.9%	111.0%	113.6%	96.5%	115.1%	110.3%	94.3%	100.2%	91.5%	121.7%	100.9%	115.8%
	Southwest	8%	85.2%	66.5%	89.9%	84.7%	99.8%	96.4%	91.4%	113.6%	97.5%	96.0%	85.9%	100.4%	87.3%
	United King..		103.2%	58.9%	128.3%	84.6%	90.1%	82.2%	100.2%	112.8%	132.4%	91.4%	139.1%	118.6%	93.4%
	United King..	0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	Total	0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Value Added Reseller	Australia								51.0%	45.0%	64.1%	66.0%	71.9%	72.0%	67.9%
	Canada	8%	79.7%	88.3%	90.9%	88.6%	108.7%	95.5%	94.8%	89.9%	93.7%	91.1%	87.4%	90.3%	91.7%
	Central	5%	127.8%	99.4%	97.9%	93.8%	107.7%	109.1%	101.5%	116.2%	109.5%	117.6%	94.1%	109.2%	105.4%
	France		65.4%	57.8%	58.4%	77.1%	52.7%	43.4%	55.0%	115.5%	143.3%	120.4%	171.9%	134.1%	99.6%
	Germany								84.8%	74.8%	74.9%	91.1%	84.1%	78.7%	82.2%
	Germany	8%	84.7%	72.3%	101.9%	114.4%	98.2%	85.2%	101.3%	82.9%	84.1%	75.7%	81.9%	81.6%	90.8%
	Northwest	4%	113.7%	131.2%	125.7%	126.0%	106.1%	129.6%	122.1%	126.6%	128.2%	137.2%	119.0%	128.0%	128.1%
	Southwest	1%	65.0%	96.1%	97.1%	94.2%	82.4%	85.2%	90.0%	95.0%	84.0%	76.7%	114.4%	90.8%	95.0%
	Southwest	2%	121.7%	104.3%	104.5%	104.4%	103.7%	112.8%	106.3%	107.5%	99.1%	105.5%	102.1%	103.7%	106.3%
	United King..				82.3%	81.4%	99.9%	80.4%	84.2%	93.8%	100.2%	94.8%	105.5%	98.1%	86.1%
	Total	0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Warehouse	Australia								177.9%	183.9%	115.8%	126.8%	82.7%	97.0%	107.1%
	Canada	5%	101.9%	91.2%	94.1%	98.1%	109.7%	99.7%	99.3%	94.1%	97.8%	102.3%	86.3%	95.4%	98.0%
	Central	3%	114.5%	120.1%	79.5%	84.9%	88.8%	94.5%	86.3%	81.4%	91.7%	80.3%	94.0%	86.2%	99.5%
	France		58.2%	50.4%	123.5%	109.8%	99.9%	102.6%	110.5%	103.5%	115.5%	100.1%	111.0%	107.8%	99.6%
	Germany								102.9%	106.3%	102.3%	82.6%	100.2%	86.8%	93.5%
	Germany	9%	77.9%	64.1%	100.1%	105.3%	109.8%	90.1%	100.3%	93.1%	79.5%	76.7%	89.5%	85.7%	86.4%
	Northwest	9%	141.6%	152.0%	85.8%	98.6%	97.5%	94.2%	93.2%	111.0%	112.6%	114.1%	109.3%	111.6%	111.4%
	Southwest	5%	127.7%	123.9%	88.2%	95.4%	69.9%	81.9%	84.0%	81.8%	77.9%	79.6%	103.6%	84.7%	107.4%
	Southwest	2%	80.5%	88.7%	109.7%	100.6%	102.1%	115.4%	108.0%	104.9%	102.6%	114.7%	90.6%	103.8%	102.2%
	United King..		69.8%	60.5%	103.3%	101.9%	83.1%	78.5%	90.9%	110.6%	108.3%	100.0%	140.8%	112.8%	96.2%
	Total	0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Grand Total		0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

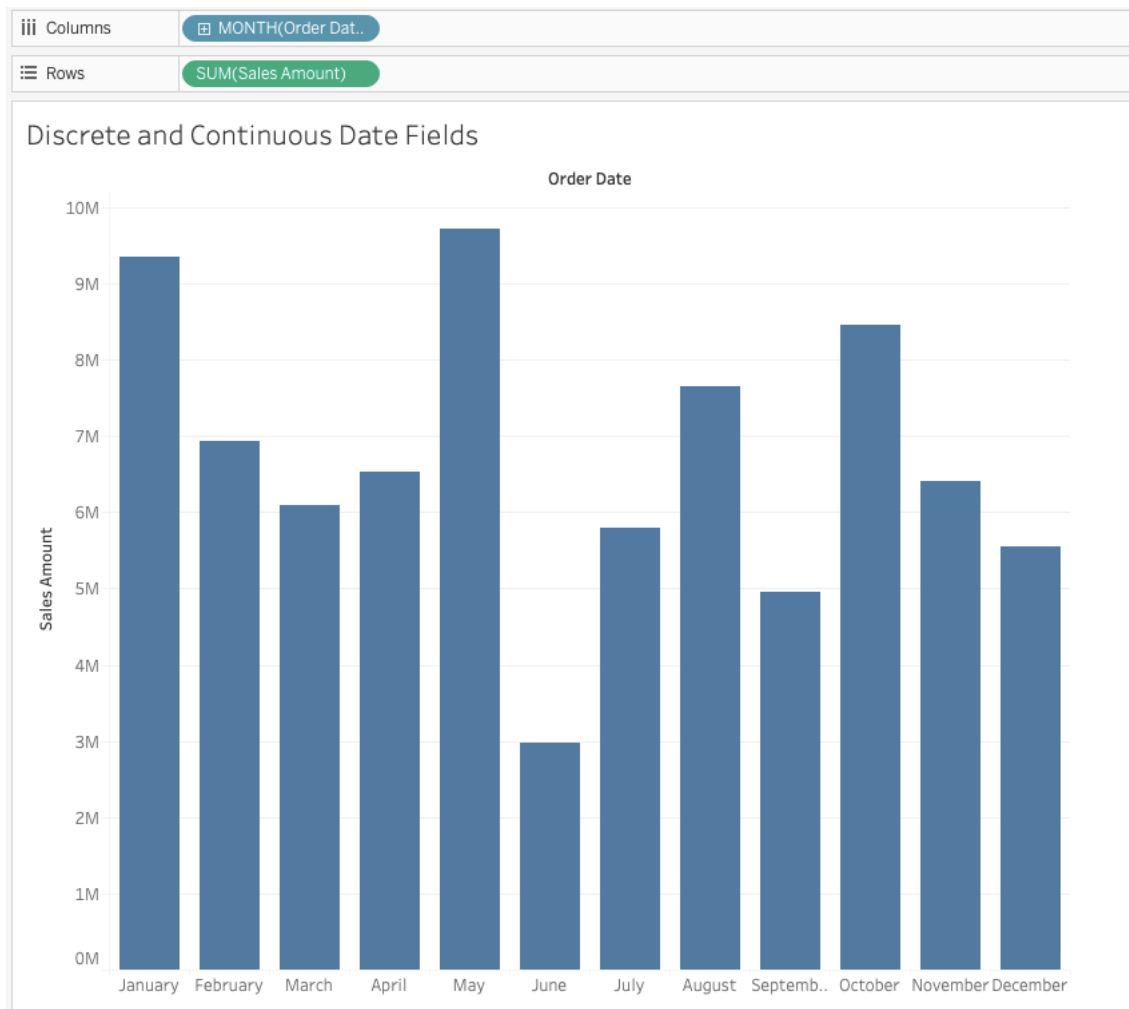
(4). Based on the cross table from (3), Why do we have **100%** as the subtotal value at each pane level?

Answer: Each pane is representing "Business Type" attribute for each quarter. As we are computing based on each pane in the downward direction, hence the subtotal is calculated as 100% at each pane level.

2. Discrete and Continuous Date Fields:

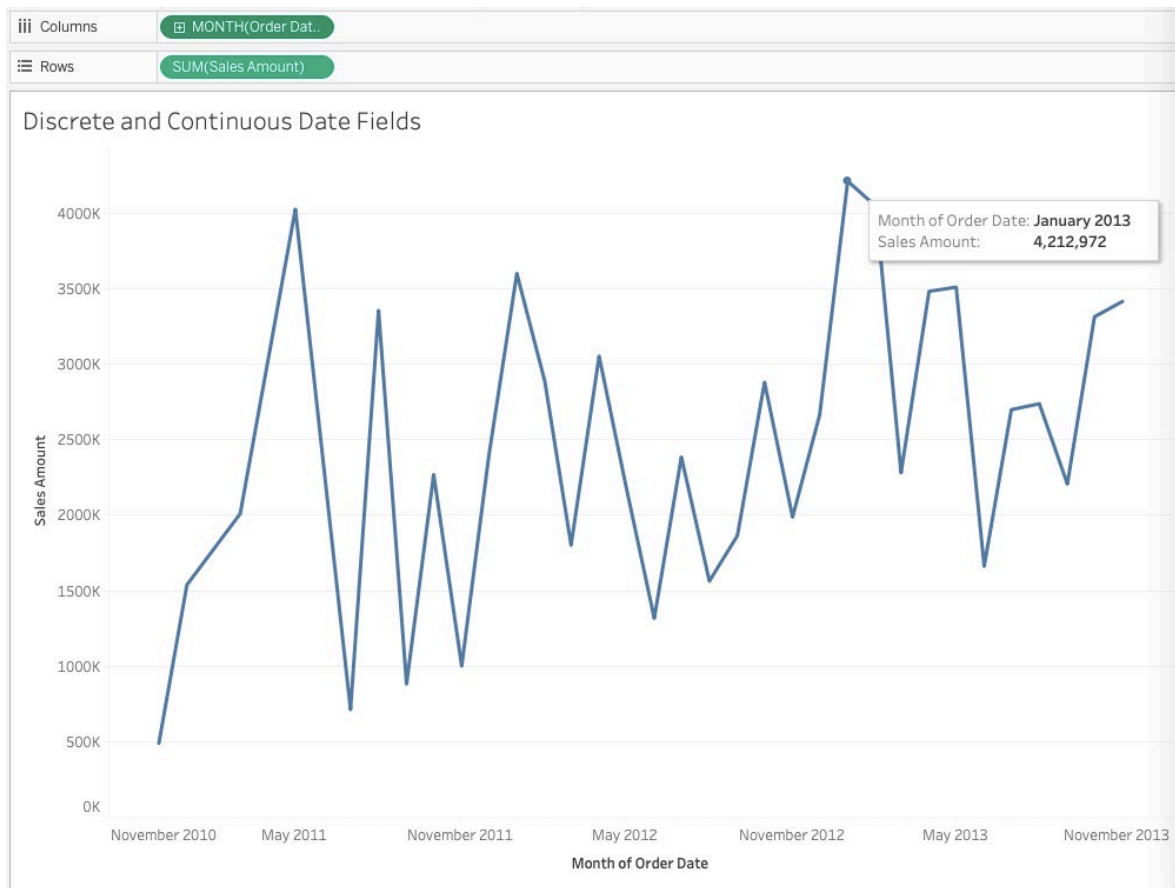
(1). Create and display a bar chart, with **sum of "Sales amount"** as rows, and **discrete type month of "Order Date"** from FactResellerSales table. Which month overall has the lowest **sum of sales amount**?

Answer: **June** month has overall the lowest sum of sales amount



(2). Create and display a line chart, with **sum of "Sales Amount"** as rows, and **continuous type month of "Order Date"** from FactResellerSales table. Which month and year has the highest sum of sales amount?

Answer: *January of 2013 has the highest sum of sales amount.*



(3). When do we usually use continuous date field, and when do we use discrete date field?

Answer: *Discrete means individually separate and Continuous means unbroken, without interruption. Discrete date field is used when my analysis requires to have distinct marks that can be sorted, whereas Continuous date field is used to look at a trend over a continuous time period. Continuous date fields cannot be sorted. Discrete data will generate rows and columns, whereas continuous data generates axes.*

3. Dual axis and Combined axis chart:

(1). Compare the **average of "Sick Leave Hours"** (from DimEmployee) and the **average of "Vacation Hours"** (from DimEmployee) using a dual axis chart, with **"Department Name"** (from DimEmployee) as columns; synchronize the two axis and display the chart;

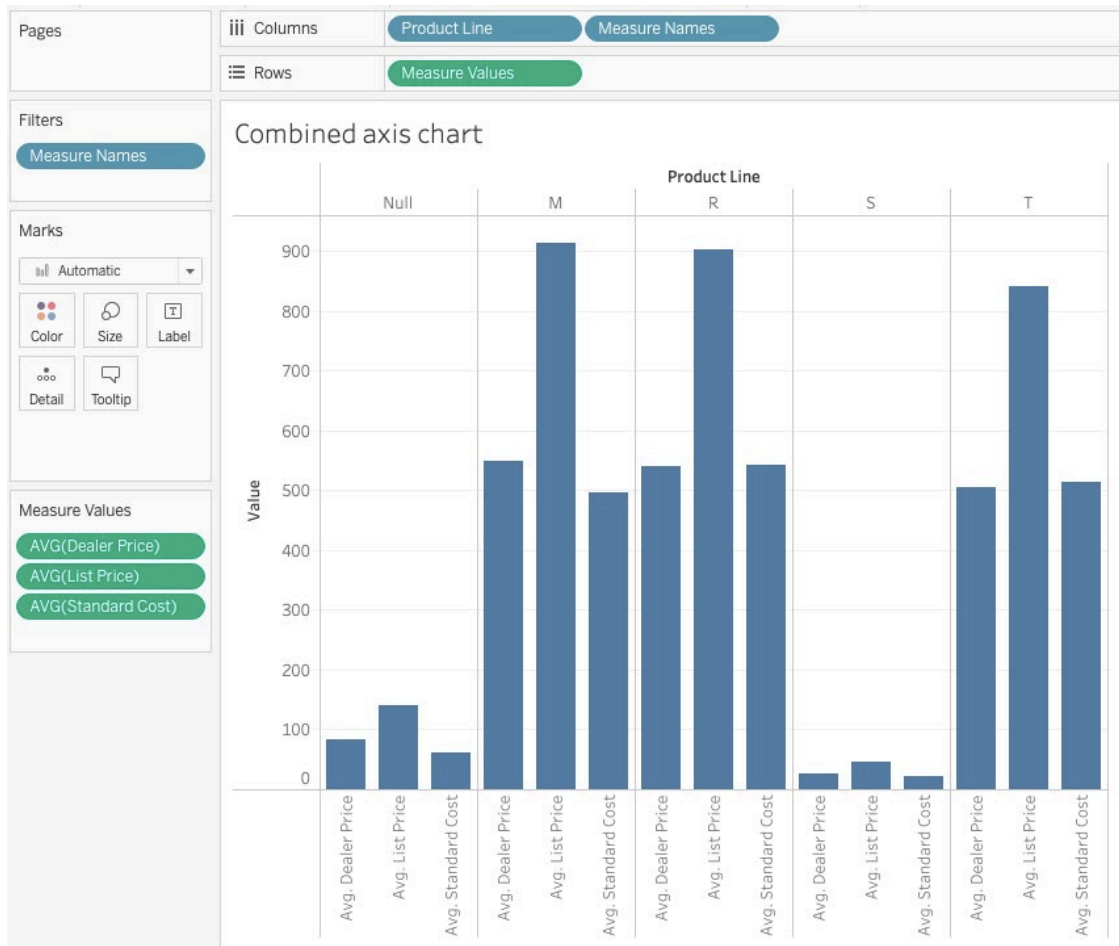


(2). Based on the dual axis chart, which departments have higher average sick leave hours than average vacation hours?

Answer: Departments which have higher average sick leave hours than average vacation hours are: -

- Engineering
- Executive
- Research and Development
- Sales
- Tool Design

(3). Generate and display a combined axis chart, with **“Product Line”** (from DimProduct) as columns, and **average of “Dealer Price”, average of “List Price”, and average of “Standard Cost”** (all from DimProduct) as rows;

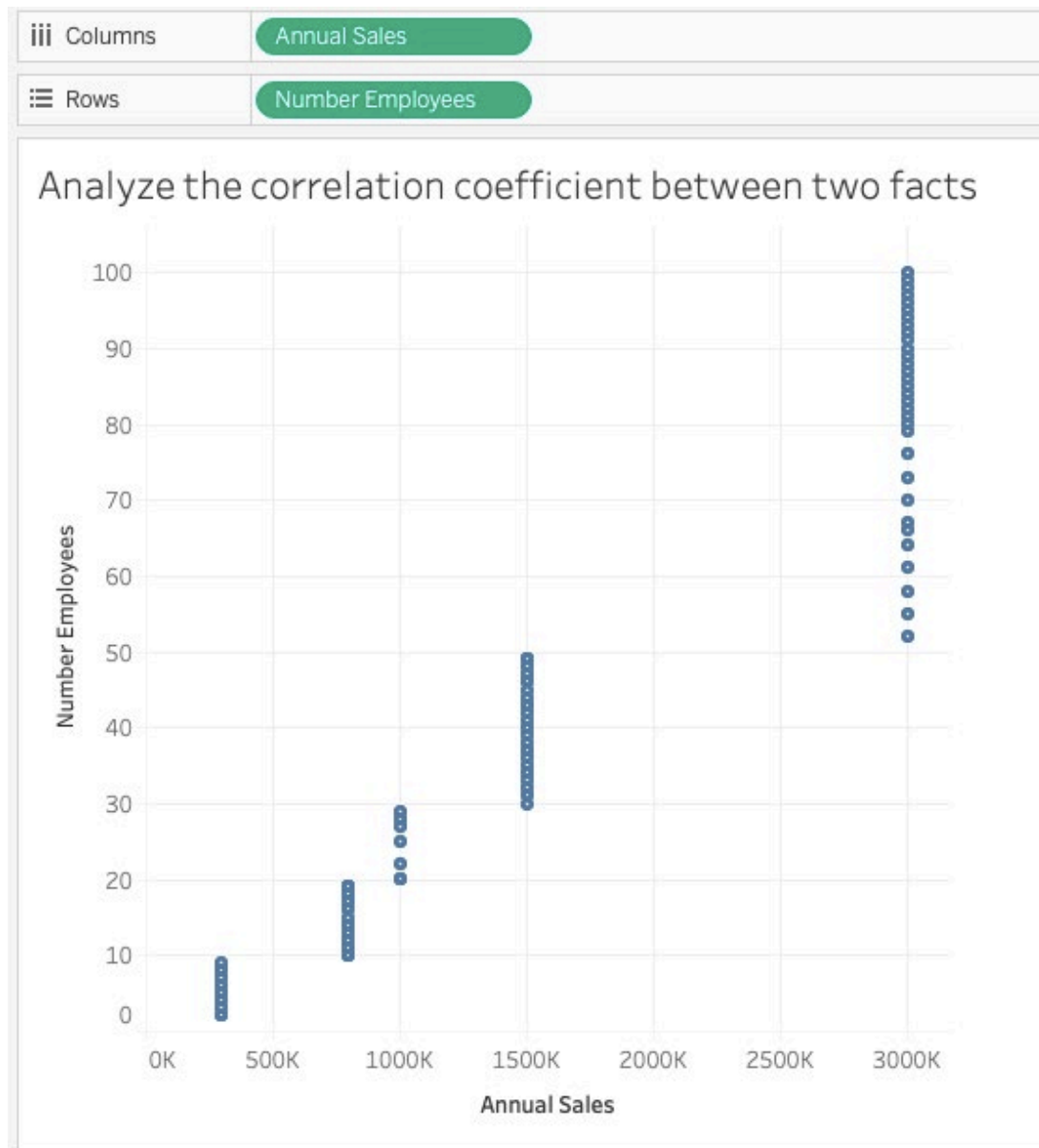


(4). Based on the prior combined axis chart, which product line has the highest average “List Price”, and which product line has the lowest average “List Price”?

Answer: The product line **M** has the highest average “List Price” and the product line **S** has the lowest average “List Price”.

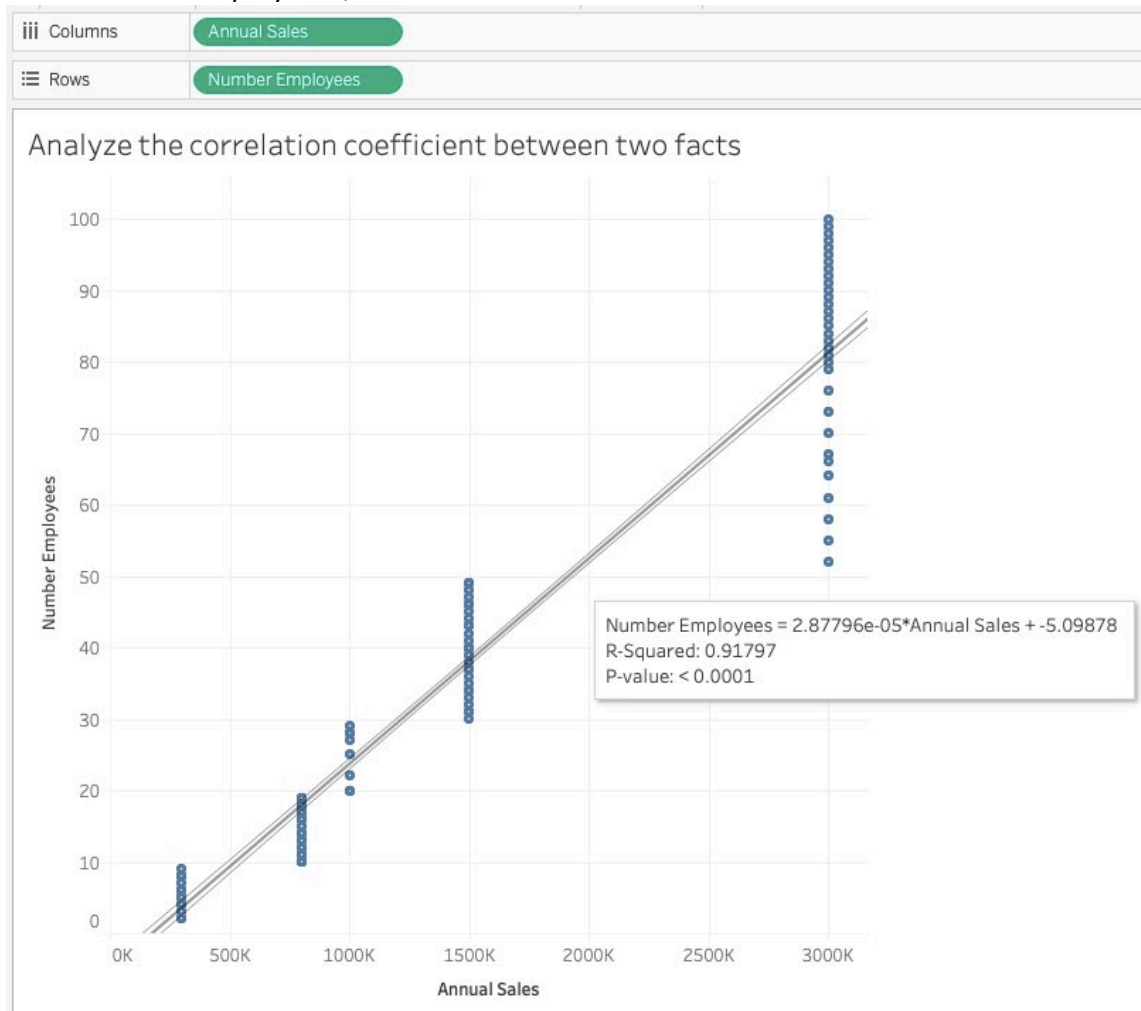
4. Analyze the correlation coefficient between two facts:

(1). Generate and display a scatter plot with facts ***“Annual Sales”*** and ***“Number Employees”*** from DimReseller; (hint: uncheck the aggregated measures)



Marks count - 701

(2). Generate and display **a linear trend line** on top of the scatter plot with facts “Annual Sales” and “Number Employees”, enable the confidence interval of the linear line.



(3). Based on the linear trend line, what is the R square value? What is the p value? What is the direction of the correlation coefficient between Annual Sales” and “Number Employees”? Are the correlation coefficient significant?

Answer: Base on the linear trend line: -

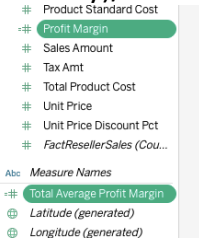
- R-squared value: **0.91797**
- P-value: **< 0.0001**
- The direction of the correlation coefficient between Annual Sales and Number Employees: **upward or positively linear**
- As the p-value is less than 0.5, the correlation coefficient between the two **is significant**.

5. Create Calculated Fields:

(1). Create and display calculated fields:

a. **Profit Margin** = ((Unit Price – Product Standard Cost)*Order Quantity)/Sales Amount;

b. **Total Average Profit Margin** = Average ((Unit Price – Product Standard Cost)*Order Quantity)/ Average (Sales Amount);

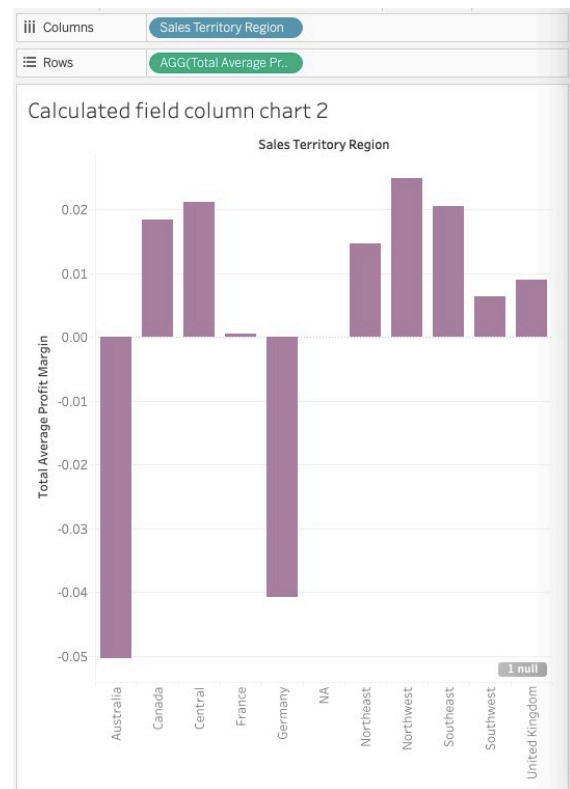
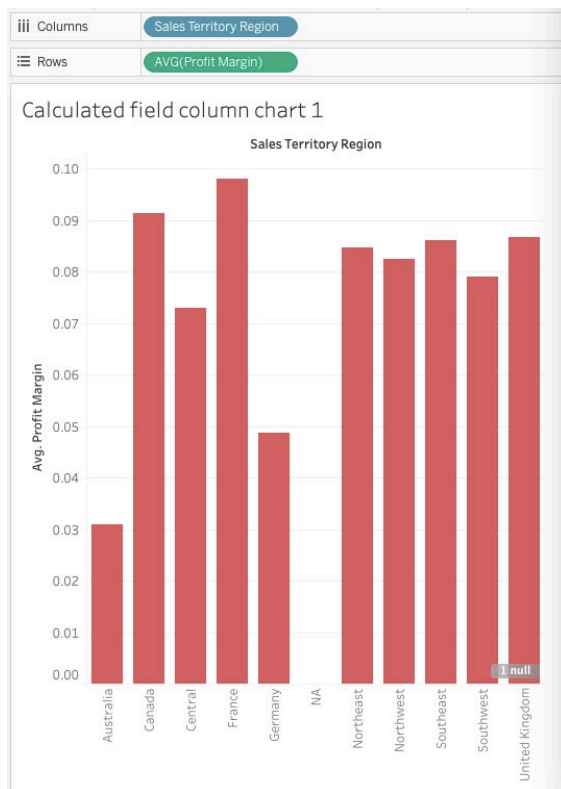


(2). Generate and display two column charts, with **average of “Profit Margin”** and **“Sales Territory Region”** (from DimSalesTerritory) in one chart, and **“Total Average Profit Margin”** and **“Sales Territory Region”** in another chart; Why are these two charts look different?

Chart 1: -

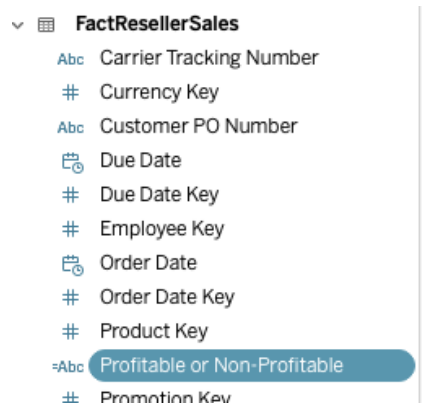
Average (Profit Margin = (((Unit Price] - [Product Standard Cost]) * [Order Quantity]) / [Sales Amount]) for a region; First one is the average of all the profit margins, where profit margin is profit for each divided by sales amount of each.

Chart 2: - (Total Profit Margin = AVG(((Unit Price] - [Product Standard Cost]) * [Order Quantity]) / AVG([Sales Amount])) for a region; Second one is average of Profits divided by average of sales amount.



(3). Create another calculated field using IF THEN ELSE statement:

Create and display a new calculated field **“Profitable”**: If “Unit Price” – “Product Standard Cost” is greater than 25, then we assign a value “Profitable”; otherwise, we assign the value “Nonprofitable”;



(4). Create and display a cross table, with the levels from **“Profitable”** as rows, and **average of “Profit margin”** as cell values in this cross table;

The image shows a screenshot of the Tableau interface. The Rows shelf contains 'Profitable and Non-Pr..'. The Marks shelf contains 'AVG(Profit Mar..'. The view displays a table with two rows: 'Non-Profitable' and 'Profitable', and one column: 'IF THEN ELSE calculated field'.

IF THEN ELSE calculated field	
Profitable and Non-Profitable	
Non-Profitable	0.07130
Profitable	0.12615

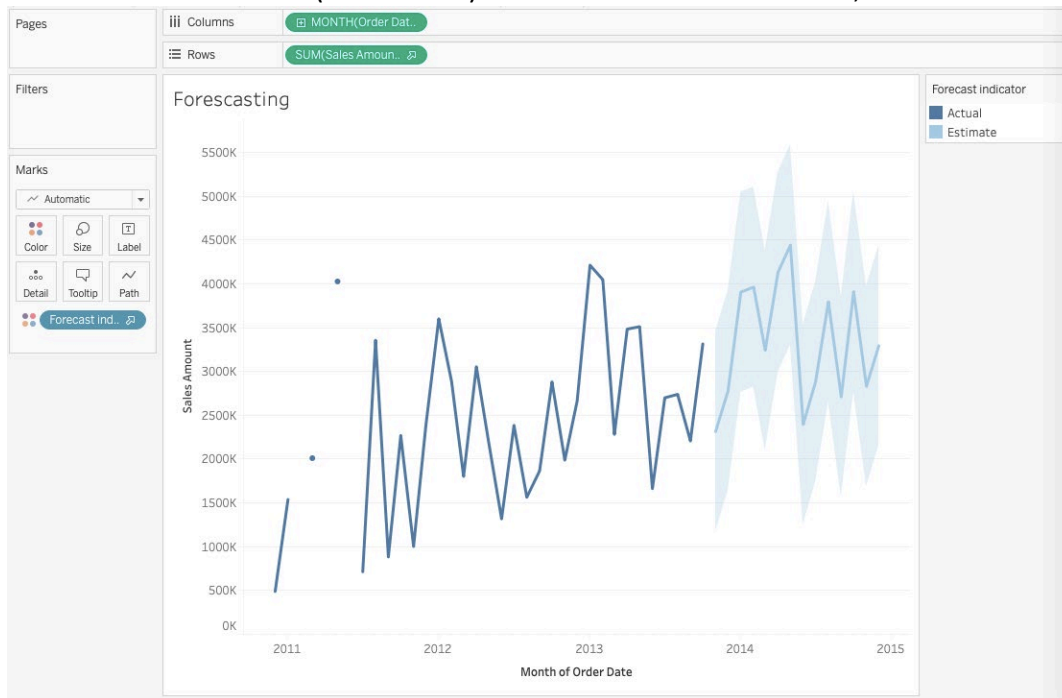
(5). What are the **average profit margins** for nonprofitable and profitable categories respectively?

Answer: The average profit margins for respective categories are as follows: -

- Profitable: **0.12615**
- Non-Profitable: **0.07130**

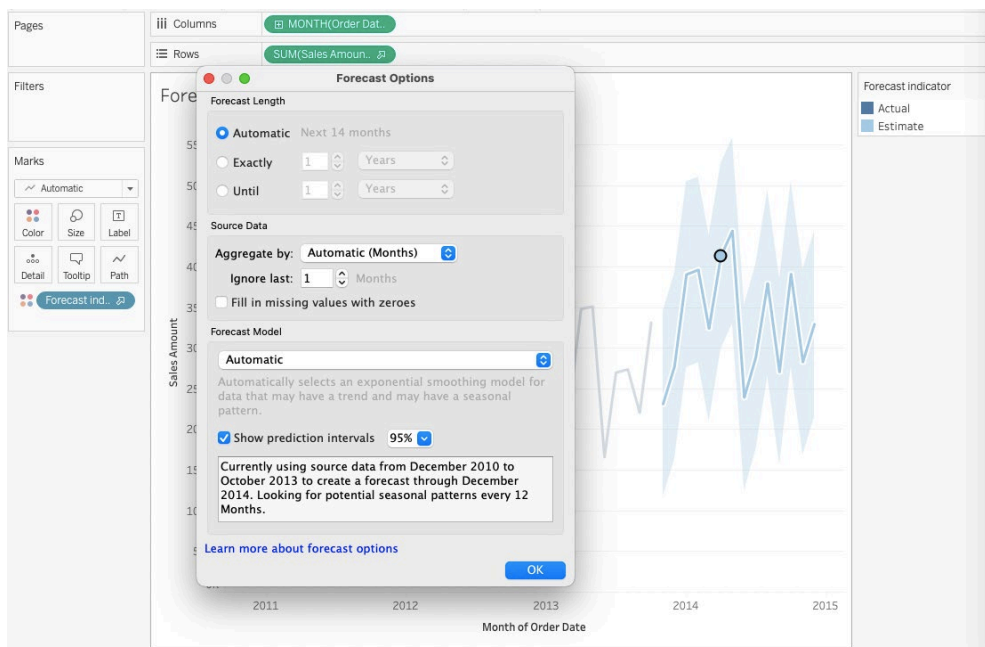
6. Forecasting and Clustering:

(1). Generate and display a forecasting line predicting sum of “Sales Amount” on the basis of month of “Order Date” (continuous) from FactResellerSales table;



(2). Use default settings from forecast options, how many months’ sum of sales are predicted in the model?

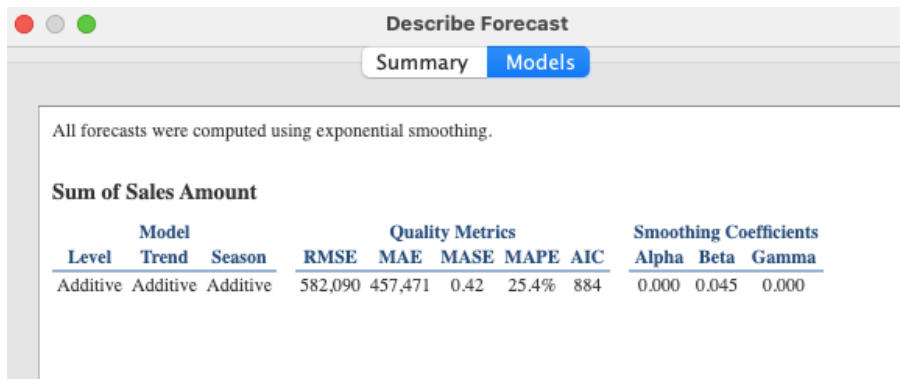
Answer: Sum of sales have been predicted for next **14 months** in the forecasting model.



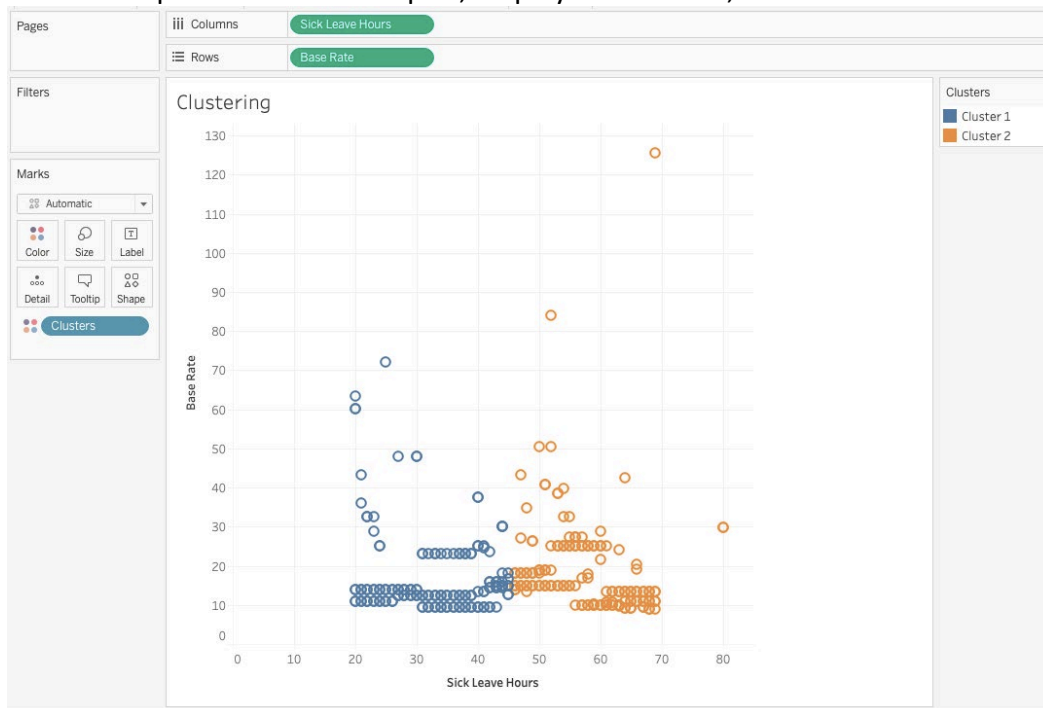
(3). From the summary of the models, what are the quality metrics of the forecasting model?

Quality Metrics				
RMSE	MAE	MASE	MAPE	AIC
582,090	457,471	0.42	25.4%	884

Answer: The quality metrics of the forecasting model are: -



(4). Generate a scatter plot with “Sick Leave Hours” and “Base Rate” from Dimemployee; Cluster the plots in the scatter plot; Display the clusters;



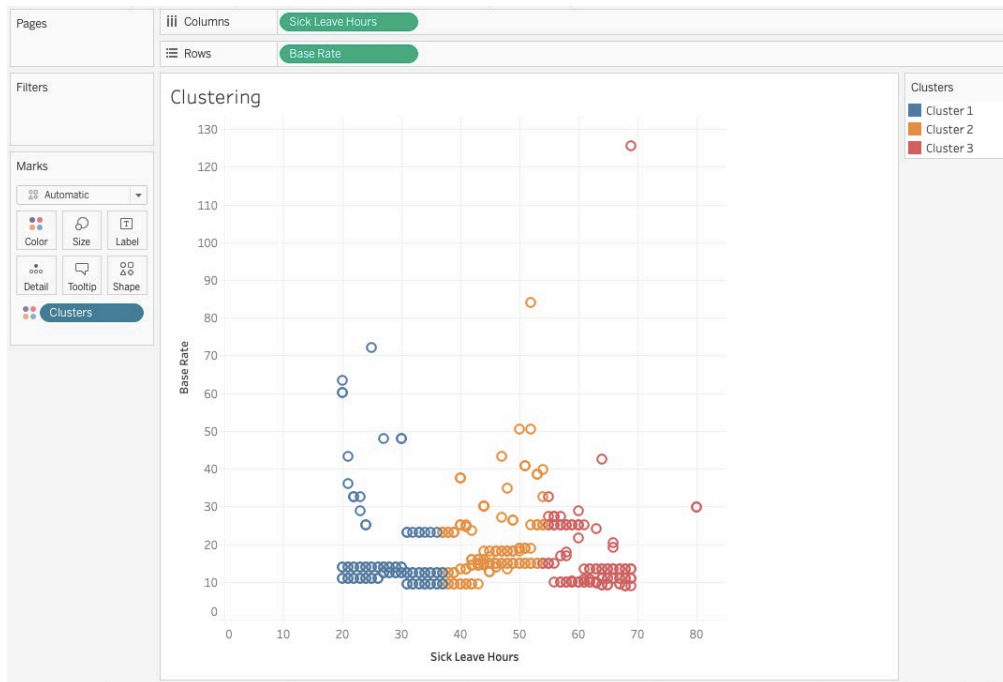
(5). Use the default setting, how many clusters are generated?

Answer: By default, two clusters have been generated.

(6). Change the cluster number to 3, display the clustered scatter plot, how many items are in each cluster?

Answer: Number of items in each cluster are as follows: -

- Cluster 1: **95** items
- Cluster 2: **106** items
- Cluster 3: **95** items



Describe Clusters

Summary Models

Inputs for Clustering

Variables: Sum of Base Rate
Sum of Sick Leave Hours

Level of Detail: Not Aggregated

Scaling: Normalized

Summary Diagnostics

Number of Clusters: 3
Number of Points: 296
Between-group Sum of Squares: 15.429
Within-group Sum of Squares: 5.4703
Total Sum of Squares: 20.899

Clusters	Number of Items	Sum of Base Rate	Sum of Sick Leave Hours
Cluster 1	95	17.998	28.095
Cluster 2	106	20.44	45.547
Cluster 3	95	16.859	62.232
Not Clustered	0		

☐ Show scaled centers

Copy to Clipboard [Learn more about the cluster summary statistics](#) Close